

ELECTRICITY INVESTMENT AND SECURITY OF SUPPLY IN LIBERALIZED ELECTRICITY SYSTEMS

**In W. Mielczarski (ed.), 2005, *Development of electricity markets*,
Series: The European power supply industry**

Richard MEADE[∞]

New Zealand Institute for the Study of Competition and Regulation Inc.
Cognitus Advisory Services Limited
Wellington (New Zealand)

Abstract

Eliciting generation investment by decentralized, profit-seeking private investors is a key goal of electricity liberalization. Debate rages regarding the ability of energy-only electricity markets to ensure that such investors provide generation investment as and when needed to ensure “the lights stay on.” Many argue that despite theoretical predictions to the contrary, energy-only markets will under-provide the requisite level of investment, due to market imperfections that are either inherent (such as consumer resistance to real-time pricing) or imposed (such as price caps to curtail market power). The nature of these imperfections is increasingly being debated, with security of supply formerly being regarded as a public good, but later analysis showing this is not the case (or even if it were, why that need not necessitate intervention). Greater attention is now being paid to externalities associated with the provision of security of supply, but evidence on the importance of such externalities is yet to be presented. Similarly lacking is evidence on the superiority of mechanisms often proposed or implemented to encourage investment in generation capacity where energy-only markets are thought to elicit inadequate investment. These mechanisms include capacity payments, capacity obligations, options-based capacity schemes and capacity subscriptions with load-limiting fuses. While the latter are argued to represent an elegant and non-distortionary means to encourage market-based security of supply, the other alternatives are shown to be conditionally optimal at best, and in principle and practice subject to self-defeating features that can be bettered by refinements to energy-only market arrangements (greater demand-side responsiveness) and structural measures (vertical integration of generation and energy retailing). By instead pursuing these alternative measures security of supply is more easily achieved, electricity prices are less vulnerable to exploitation of generator market power, and generation investment is more likely to arise. The need for price caps, which then necessitate compensatory capacity mechanisms to elicit investment, is then reduced. At the same time exposure to regulatory risk is lessened. Combining these measures with greater political and regulatory restraint is argued to provide a more stable and superior means to elicit the investment needed to provide the socially optimal security of supply, addressing any market imperfections at source rather than introducing new mechanisms at least as much at risk of imperfection. The use of capacity mechanisms is argued to increase the risk that energy-only markets will fail to perform as expected and required, undermining the liberalisation process. As such they raise the prospect that governments and regulators concerned about security of supply will once again find themselves responsible for achieving it, at consumers’ and/or taxpayers’ expense, but with lesser prospect of success.

1. INTRODUCTION

How can the security of electricity supply best be ensured, with ever-growing electricity demand, in a liberalized, competitive electricity market? This is a fundamental question

[∞] Helpful comments from Glenn Boyle and Lew Evans are gratefully acknowledged, as is the research assistance of Lisa Ryan. All errors or omissions remain the responsibility of the author, who can be contacted at: richard.meade@cognitus.co.nz.

receiving increasing attention by electricity reformers and policymakers worldwide, particularly following spectacular recent supply interruptions in North America (California 2000/01; North-Eastern US and Ontario 2003) and Europe (London, Denmark/Sweden and Italy in 2003) – regardless of whether these interruptions were due to features of reformed electricity systems or simply to transmission vulnerabilities. It is naturally a question of political importance since lawmakers face significant discomfort if market reforms they have supported either experience supply difficulties or simply appear to be less secure than pre-reform systems. For those contemplating liberalization – which inevitably involves a step into the unknown, not least because it places the burden of new investment decisions on decentralized, unbundled private investors rather than centralized state or private integrated monopolies – they need to be reassured that this will not result in “the lights going off.”

This fundamental question is naturally transformed into subsidiary questions also vexing the minds of reformers, particularly as any excess capacity inherited from pre-reform systems begins to dwindle. These include, how much investment is required to ensure supply security in a liberalized system, and how can this investment, by private parties, be encouraged in a timely and efficient fashion? Both questions are predicated on a presumption: that liberalized electricity markets cannot, of themselves, deliver the required investment in system capacity, whether in transmission, generation, or demand-side flexibilities, due to some form of “market failure.” Alternatively, it is presumed that there are features of liberalized electricity systems imposed by reformers, such as wholesale price caps to address concerns about market power, which hamper the necessary investments that would otherwise occur. Either way there are even more fundamental questions implicit in these subsidiary questions, namely, what exactly is supply security, and should reformers even be concerned about trying to achieve it?

As yet there is little consensus on these questions, insofar as they are even being addressed. Many reformed or reforming electricity systems, notably in Europe, include no explicit features to ensure security of supply, with the England and Wales system even abandoning such features when its pool was replaced by NETA in 2001. Others, such as Argentina and Spain, and PJM, New York and New England in the US, have implemented a variety of mechanisms such as capacity payments or capacity requirements, with the explicit intent of improving the incentives for private investment so as to ensure supply security (with or without otherwise conflicting impositions). Given the differences between reformed systems, which obscure comparisons of approach, and the relative infancy of functional examples of market arrangements to encourage the large, sunk and long-lived investments typical of electricity systems, the jury remains out.

This paper contributes to the current debate by deconstructing both the questions that investment-encouraging mechanisms such as capacity markets attempt to answer, and the answers that they provide. It argues that where market imperfections are thought to hinder necessary investment, and the term “necessary” in this context is used guardedly, the introduction of such arrangements are likely to be inferior to instead addressing the causes of those imperfections. This inferiority stems not only from the fact that these arrangements exacerbate any departures from the “ideals” of market-based reforms, but also because of the inevitable costs associated with imperfect interventions.

Additionally, two structural features of reformed electricity systems, one common (oligopolistic generation) and one less common (generators vertically integrated with energy retailers), offer overlooked advantages in terms of encouraging investment. Consequently, arrangements commonly introduced to encourage desired investment, or to hinder perceived evils of market power, will be self-defeating, or at least frustrate the achievement of any desired supply security objectives. Thus it is concluded that the focus of reformers should more productively be directed towards addressing any undue and/or fundamental obstacles to electricity investment, helping to better address multiple reform aims simultaneously, while acknowledging that this approach would require both continued boldness and self-restraint on the part of reform sponsors.

The paper is organised as follows. Section 2 addresses the last-stated fundamental question outlined above – what does supply security mean in a liberalized electricity market?

Sections 3, 4 and 5 address the second and third fundamental questions – what, therefore, is the “required” level of new investment, and how can timely electricity investments by private parties be encouraged to provide any requisite level of supply security? Section 3 reviews the theoretical conclusion that a freely-operating energy-only electricity market should be sufficient to elicit the socially optimal level of investment, while Section 4 surveys and critiques the reasons typically offered for why this conclusion might fail in practice, and why, therefore, interventions may be required to ensure optimal investment. Section 5 reviews and appraises some of the most common interventions suggested to remedy supposed market failures argued to lead to underinvestment. In the light of these critiques and appraisals, Section 6 then offers suggestions for answering the final fundamental question, posed at the outset – how can supply security best be ensured? Section 7 concludes.

2. “SECURITY OF SUPPLY” IN LIBERALIZED ELECTRICITY MARKETS

Hirst and Hadley (1999), and Oren (2000), unpack the North American Electric Reliability Council concept of “reliability” and its two component parts, “security” and “adequacy”. Reliability is defined as “the degree to which the performance of the elements of [the electrical] system results in power being delivered to consumers within accepted standards and in the amount desired.” Adequacy, in turn, refers to the subject of this paper (security of supply), meaning “the ability of the system to supply the aggregate electric power and energy requirements of the consumers at all times,” which has a timing focus from the immediate to the longer-term. Security refers to a more narrow version of adequacy, being “the ability of the system to withstand sudden disturbances,” i.e. contingencies such as generator outages or grid failures, which takes on a more real-time and operational focus given the capacity for such disturbances to create outages spread across interconnected AC grids. Adequacy is influenced by the timing and level of investment in generation and transmission capacity, but also by demand-side measures such as provisions for interruptible load. Clearly a system that has significant capacity (or interruptible load) should be more secure, all other things being equal. Conversely, even a system lacking capacity might be operated in a stable manner.

International Energy Agency (2002, p.9) states that “In the long term, the security of electricity supply depends on the adequacy of investment in terms of providing: enough generating capacity to meet demand . . .” as well as protections against variations in fuel supply, and adequate grid and distribution networks for energy transportation. Supply security is also taken to refer to the likelihood that energy will be supplied without disruptions, and the influence of reliability on electricity price levels and volatility, but not the reverse, is noted. In New Zealand security of supply is taken to require that: “Sufficient generation capacity is built or energy efficiency improvements made to meet ongoing demand growth” and that “The system has sufficient reserve energy (plant and fuel, or contracted demand response)” to deal with adverse hydro situations or unexpected supply disruptions, as well as transmission/distribution and capacity/fuel requirements similar to the IEA definition (Ministry of Economic Development (2004)).

From an economic perspective all three definitions are wanting in the context of liberalized electricity markets. All three contain a presumption, if not explicit statement, that electricity systems should meet all demand for electricity, always, irrespective of the cost of achieving this, and with no reference to the role of electricity prices and markets in equating available supply and demand. Aside from the technical impossibility of this aim, in so doing all three definitions betray a fundamental ignorance of the role and function of electricity markets, and/or a lingering resilience of the discordant engineering view of electricity systems that pervaded electricity planning prior to liberalization.

Prior to the introduction of electricity markets, during which time electricity prices were determined administratively and at a level ensuring a balance of supply and demand only by chance, growth in electricity demand was viewed as an exogenously determined quantity, and it was the task of engineers to build electricity systems as efficiently as they could to

ensure that demand growth was satisfied. Where demand could not be satisfied, whether due to contingent events, insufficient funds for investments, or simply poor planning and delivery, electricity demand was temporarily endogenised through administrative rationing, such as via blackouts, with no knowledge of or regard to different consumers' preferences for ongoing supply (save, perhaps, for "essential services" such as hospitals). Where demand was always met, the experience worldwide was that this was often a consequence of expensive overinvestment in capacity, of the ubiquitous "gold-plated" electricity systems, funded by consumers and/or taxpayers for whom electricity prices were levied as if taxes (for example, see Roques et al. (2004) re Britain under the CEGB).

Following the introduction of electricity markets, to varying degrees demand is now a more endogenous component of the electricity system, interfacing with supply to simultaneously determine prices at which the two are in constant balance. Price and demand are no longer determined wholly outside of the electricity system, but critical co-determiners of the system's operation and progress.¹ Electricity system planning is thus no longer the preserve of engineers and optimal control models, not to mention politicians, but shared with consumers to the extent their preferences and actions influence the evolution of market prices. These prices in turn provide critical signals to private parties now expected to determine when, where and how to profitably invest in new capacity, and to bear the risks of poor decisions and adverse events subsequently affecting their investments.

Liberalization has also fundamentally shifted the risks of electricity investment: "One of the theoretical major benefits of liberalization lies indeed in the redistribution of risks among the different stakeholders of the electricity industry" (Roques et al. (2004, p. 31)). Overinvestment is no longer underwritten at consumer and/or taxpayer expense, whether it arose due to engineering conservatism or political/regulatory aversion to backlashes following blackouts. Instead, now both investors and consumers bear the risk of under or overinvestment, with electricity prices being an important adjudicator of where the burden of those risks will fall. To the extent that politicians and regulators continue to influence the shape and progress of electricity markets, however, they too must share responsibility, to both investors and consumers, for their impact on market arrangements and hence prices.

Market-determined electricity prices serve other critical functions that conflict with the presumption contained in the above three definitions of security of supply/adequacy. So long as prices are determined with at least some demand-side bidding, they reflect consumers' evaluations, however imperfectly, of the importance of having electricity supplied. Since prices reflect the evaluation of the "marginal" consumer, it can then be presumed that all consumers not willing to pay that market price do not value electricity supply at that time as highly. Their demand is rationed off by the market price. The converse, that those who are supplied at the prevailing market price value ongoing supply at least as highly as the marginal consumer, may or may not be true. Since many consumers continue to be supplied on fixed price supply contracts they do not face changing prices as a means to ration their consumption in response to changing supply/demand balance, and hence it might be thought that their willingness to pay for supply is not being gauged. However, since many of those consumers who are so exposed to changing prices choose to hedge their price risk, the price at which they are prepared to hedge that risk provides some measure of their willingness to pay for ongoing supply.² That the value of not being supplied varies across consumer types,

¹ Although you might be forgiven for thinking this is still not so despite decades of liberalisation, based on official forecasts of New Zealand electricity demand growth. Such estimates are used as a basis of estimating "required" levels of new generation, but ignore the role of electricity prices in rationing demand (see, for example, Ministry of Economic Development (2003)). Unless demand growth forecasts are derived in association with simultaneous forecasts of market-clearing electricity prices, it is impossible to infer what level of generation investment is required to serve that demand at that price.

² Hedging against electricity price risks should be considered a rational insurance choice by those consumers who suffer great cost from price volatility, and/or who have limited capacity to mitigate that risk in other ways (e.g. via investment in energy efficient or flexible technologies). For smaller electricity customers, for whom electricity bills represent a small fraction of their weekly household expenditure, fixed-price supply contracts can represent a relatively cost-effective means to mitigate their electricity

and as a function of the length of supply interruptions, is well documented (see, for example Roques et al. (2004), de Nooij et al. (2004)). Thus the presumption that all consumers desire ongoing supply, in all states of nature, at all times, and at any price, is clearly wrong. Where market prices continuously adjust to ensure that available supply, however small, is balanced with demand, however potentially large (which is now a question of price), the presumption is also redundant.

Finally, the above definitions of security of supply/adequacy contain at least two other conceptual flaws. First is the failure to recognise that generation and transmission capacity are not the only means to achieve adequacy. Where these definitions emphasise such supply-side features they again betray the traditional engineering and central planning focus. The New Zealand definition at least acknowledges the role played by demand-side features such as energy efficiency and interruptible load in helping to reduce demand rather than increase supply capacity to achieve adequacy. However, even that definition neglects the potential role of providing consumers with options and incentives to vary their consumption in response to changing electricity prices in also achieving that aim. Second is the presumption that rationing should always be avoided. Joskow and Tirole (2004) show that some rationing may in fact be socially optimal as a means of proxying for demand-side price-responsiveness where it is not otherwise available. Oren (2000, p. 3) asks “What is the meaning of lost load in a competitive market with no obligation to serve?” Hirst and Hadley (1999) similarly exclaim “Is adequacy even a relevant term for a restructured electricity industry?” Perhaps the concept remains meaningful if electricity markets are not effectively rationing electricity to those prepared to pay for it at all relevant times. But if they are, supply security boils down to asking little more than “who is prepared to pay what to obtain electricity supply at some probability that that supply will occur?” In that respect it is little different to the security of supply of cabbages.

3. OPTIMAL INVESTMENT UNDER ENERGY-ONLY PRICING

Before electricity liberalization, when electricity investment planning was centralized and coordinated under the control of engineers and politicians, electricity supply was not free of interruptions. Electricity Shortage Review Committee (1992) documents decades of forced rolling blackouts in New Zealand arising under such an electricity system. To make matters worse, Galvin (1985) shows that these arrangements also failed to deliver electricity at least practicable cost – the objective of the day. In fact over-investment occurred due to systematic and gross over-estimates of demand growth, and investments suffered commissioning delays, large cost overruns, and electricity production costs well in excess of those predicted. Conversely, Evans and Meade (2005) record how tight supply situations caused by adverse hydrological events post-liberalization in New Zealand have not required compulsory rationing. The non-market based and administrative approach to electricity planning, distorted as it was by political price-setting and investment decisions, resulted in both unreliability (e.g. blackouts) and over-expensive capacity. Such considerations are worth bearing in mind when comparing different approaches to securing the investment required to satisfy ongoing demand for electricity.

From the preceding discussion it should be clear that any statement specifying a level of generation investment that is “required” in order to ensure security of supply contains too many degrees of freedom (e.g. at what price, and in what state of nature?). It also contains a strong assumption regarding the distribution of consumer preferences for ongoing supply (i.e. an estimated measure of the marginal consumer’s preference for ongoing supply). Jaffe and Felder (1996), by contrast, present a static economic analysis for determining the optimal level of generating reserve margin (which reasonably presumes at least some consumers value ongoing supply), from which a “required” optimal level of generation investment to

price exposure. For larger customers, such contracts can provide useful price insurance in the same way that fixed-rate mortgages do for borrowers unwilling or unable to bear interest rate risk.

achieve that margin might be inferred. Following standard economic argument, the optimal reserve level arises where the marginal social benefit of avoiding unmet demand just equals the marginal social cost of supplying that load. In practice the former is typically measured using estimates of the value of lost load, VOLL, multiplied by an estimated loss of load probability, LOLP (i.e. estimating the expected social cost of outage). The marginal cost of supply is measured in terms of direct and indirect costs of providing supply capacity, in particular, of peaking plant. However, such analysis is not directly translatable into a dynamic context, which is pertinent since generation investment is a flow concept and not a stock, as the timing of new investment and corresponding electricity price paths remain unexplained.

Dynamic simulation models of generation investment are developed by authors such as Hobbs et al. (2001), Ford (1999) and Botterud and Korpas (2004). Hobbs et al. compare alternative market schemes (discussed further in Section 5) based on an assumed, fixed reliability standard which may or may not reflect the dynamic reliability preferences of the marginal consumer over time. Hence their analysis may shed light on the timing and amount of generation under differing market schemes which meet that reliability standard, but that does not predict the “required” level of investment reflecting changing consumer reliability preferences. Ford, and Botterud and Korpas, focus on optimal investment timing and resulting electricity price paths so as to maximise investor profits, but this too does not predict a socially optimal or required flow of investment. Alternatively, Joskow and Tirole (2004) model social-welfare maximising generation investment and demand rationing under different market schemes (also discussed further below). Their analysis states conditions, such as to do with optimal price cap levels, under which social welfare is maximised, but the difficulty in measuring key parameters means it is less than straightforward to translate their results into a specific investment plan that could be said to be “required.”

In principle, so long as there is some market-determined price at which supply available at that price is sufficient to meet demand at that price, the socially optimal amount of electricity is being produced (subject to the absence of market failures, as discussed in Section 4). Thus market prices contain at least some information regarding consumer preferences for adequacy and signal these preferences to generation investors. Simultaneously they reflect the costs of supplying electricity, signalling to consumers what they need to pay for their preferences to be satisfied. Should demand grow relative to possible supply, which as a technical matter increases the risk of real-time system insecurity, the average market price will also tend to rise. As the market price, being an observable statistic providing important information about future expected electricity prices, nears the level at which the cost of new generation is expected to be profitable, generation investment will be elicited (if it doesn't, price will continue to rise until it is more than profitable to invest, making new investment even more likely). Indeed, long-term forward supply contracts may also be entered into to elicit or underwrite investments, once again with their strike prices informed by current spot and forward prices. As new investment occurs (or previously unviable plant becomes profitable at current electricity prices and is re-commissioned), energy prices would stabilise (and perhaps even fall), unless and until further new investment becomes warranted, and real-time system security should improve.³

Such a dynamic process ensures that consumers obtain the level of security over time that they are willing to pay for, given generation investors' appetites for profits and risk – the simultaneous balancing of these two considerations via the market price leading to the socially optimal solution. In practice, therefore, any level of generation investment said to be “required” which does not at the very least allow for some market-based estimate of the distribution of the value of lost load, the rationing influence of market prices, and the profitability (and riskiness) of new generation, should be viewed with scepticism. Where the party making such claims is neither a price-responsive customer, nor an investor in capacity, their information and incentives must be considered wanting.

³ Of course, if new generation or transmission technologies evolve that reduce the costs of supply then it may be possible to elicit new investment even with declining expected prices.

Theoretical support for these propositions, that energy-only markets for electricity can support socially efficient levels of investment over time, has long been known. Joskow (1976) surveys the earliest contributions, including that of Boiteux (1949) (discussed in Oren (2000), Bidwell and Henney (2004), and Rosques et al. (2004)). Later contributors include Caramanis (1982) and Chao (1983). The early theory of peak load pricing suggested uniform capacity payments to generators based on the fixed costs of peaking technology to augment energy revenues based on marginal cost. As noted earlier (Jaffe and Felder (1996)), optimal capacity arises when the marginal cost of new capacity equals the marginal cost of unserved load (i.e. marginal benefit of avoiding that unserved load), so capacity prices can therefore be approximated by the product of LOLP and VOLL. Chao (1983, p. 179) summarises the generally accepted conclusion of this theory as being “offpeak customers should simply pay the marginal operating cost and the peak period customers should bear all the capacity costs.” Oren (2000, p. 5) states “Theoretical rationale and practical experience suggest that energy-only markets with prices that are allowed to reflect scarcity rents will generate sufficient income to allow capacity cost recovery by generators.” Thus, such contributors emphasise the importance of profits made by suppliers during occasional and short-lived peak periods, so-called “scarcity rents”, in recovering capital investment costs which are not otherwise recovered during off-peak periods (in which competitive markets will price at the marginal cost of delivery). The current debate centres around whether the conditions under which these theoretical propositions hold are satisfied in practice, due to inherent obstacles or artificial impositions, and the relative efficacy of different interventions to compensate for when they do not.

4. MARKET FAILURES LEADING TO SUBOPTIMAL INVESTMENT

4.1. The Experience to Date

The reality of competitive, liberalized electricity markets is that significant investment has indeed occurred and reserve margins (measured as the excess of installed capacity over peak demand as a percentage of installed capacity) have largely been maintained. International Energy Agency (2002, Table 2, p. 30) records that generating reserve margins have remained unchanged or slightly increased in four of a selected sample of seven reformed markets, with significant declines occurring in the Australian states of Victoria and New South Wales. For the period 1985-1999 it shows (Figure 1, p. 22 and Table 1, p. 23) stable reserve margins of 30-35% in Europe, and falling but then recovering margins of almost 35% in Japan. Margins in the US, by contrast, are shown to have almost halved, from 30%, over this period. In New Zealand reserve margins are recorded as falling from 37% in 1985 to an average of just over 30% in the 1990s. Roques et al. (2004) show net investment levels falling in England and Wales with the transition from the pool to NETA in 2001, and while capacity margins increased initially under NETA to around 30%, they have since declined rapidly and are forecast to most likely undershoot the grid company’s 20% planning margin by the end of the decade. Botterud (2004) reports a significant margin of installed capacity over peak load in the Nordic power market for 1994-2003. De Luze (2003, pp. 11 and 20) paints a bleaker outlook for worldwide electricity investment, noting a “destruction of confidence in the power sector, caused by the confluence of many seemingly unrelated events” (such as the Californian and Enron meltdowns, collapsing prices due to excess capacity, and stalled deregulation), and that “the merchant model for power generation was destroyed in 2002.” He singles out regulatory issues, however, as a major reason for falls in investment levels, “Lenders and investors thought they understood the rules governing the power system when they invested. The destruction of the clear understandings investors thought they had has caused them to abandon the sector, resulting in a collapse of equity and debt prices. This disengagement will last until a new understanding of a new reality is gained and confidence is restored that this understanding will not be shattered again . . .

Regulators need to create a climate of stability in which reasonable expectations of investors can be fulfilled” (p. 6).

Even with largely enduring positive reserve margins, involving ongoing investment as well as demand growth, liberalized electricity markets have experienced significant price spikes when demand has risen and/or supply been constrained. Botterud (2004) attributes two transitory but noteworthy electricity price spikes since 2000 in the Nordic market to capacity shortage, but a longer-lived and more severe rise in prices in winter 2003 due to an energy shortage caused by adverse hydrology. New Zealand has experienced two major episodes of high winter spot power prices, mid 2001 and mid 2003, mainly due to adverse hydrological conditions, as well as many temporary spikes due to transmission outages since its wholesale market commenced operations in October 1996 (Evans and Meade (2005)). In Britain NETA faced its first stress test in 2003, with generation closures or mothballing following historically low wholesale prices in 2002 leading to falling capacity margins and steep rises in forward electricity prices (Roques et al. (2004)). The US Midwest experienced a major price spike in 1998 (Hirst and Hadley (1999)), and the Californian failures of 2000/01 are well known (Borenstein (2002)).

Electricity markets are not alone in enjoying ongoing investments in the face of such volatility in prices. Indeed, the message of the authors cited above is that such volatility, specifically the high transitory peaks in prices, are essential to fund capacity costs in competitive markets. Ford (1999) notes that construction cycles have been observed in a variety of industries, including the highly volatile agricultural and metallic commodity industries. Aluminium prices, in particular, have cycled by more than 100% over 1970-2000. He quotes (p. 638) various authors on the consequences of such volatility (re aluminium, “competition leads to both volatility and periodic price spikes, during which producers recover most of their fixed costs”), and on the implications of similar price volatility in electricity markets, including the likely chronic excess of capacity over typical demand (cf reserve margins) and the adoption of tactics by market participants to exploit high price volatility for profit. Evans and Meade (2005) report that post-liberalization investment levels have matched those pre-liberalization in New Zealand, and this despite coinciding with volatile wholesale prices and the advent of more stringent environmental legislation.⁴

It seems that widely differing views are adopted regarding the efficacy of markets for eliciting investment, particularly where energy prices are volatile and prone to spikes. One view says this is both required and consistent with theoretical prediction in order for investment to be profitable. The other views it as a sign of chaos and a justification for intervention. It is telling that after two decades of electricity liberalization in New Zealand and seven years of wholesale market operation, consulting engineers (among many others) are yet to be persuaded that the market is eliciting the right investment at the right time and at the right cost: “Development is proceeding on an ad hoc basis. We do not know which schemes are the most beneficial to New Zealand as a whole and hence we cannot be sure that the most beneficial schemes are being developed” (Sinclair Knight Merz (2003)). It is more telling that similar remarks are not commonly made about the socially optimal level of aluminium industry investment given the high volatility in its output price. The regulated and/or centralized origins of electricity industry planning, incomplete and volatile nature of the transition to liberalization, and relative newness of the technological and theoretical innovations that have paved the way to decentralization and competition are likely, partial explanators for this conceptual inertia.

4.2. General Reservations about Electricity Markets

Despite evidence of ongoing generation investment and absence of a general collapse in

⁴ Indeed, to the extent generation investment is more recently being deferred in New Zealand, a report commissioned by government indicates that this is in fact due to significantly increased regulatory risk over the past five years (PricewaterhouseCoopers (2004)).

reserve margins following liberalization, in general terms there is a concern that electricity markets are immature, their design evolving, and that electricity market participants have insufficient experience to operate them well. As such it is sometimes suggested that electricity markets cannot be trusted to provide ongoing investment so as to maintain uninterrupted supplies, at least not until arrangements have matured. It is often further suggested that interventions may thus be warranted to overcome any such market deficiencies, such as in the form of capacity payments or obligations to smooth and stimulate generation investment (Hobbs et al. (2001), de Vries and Hakvoort (2004), Botterud and Korpas (2004)).

Such concern about and distrust of electricity markets is not helped when supply difficulties first arise under new electricity markets. Purchasers, whether wholesale or on fixed tariffs, regulators and politicians are often ill-prepared for the extremes to which wholesale electricity prices must rise in order to ensure demand is rationed to available supply (with or without blackouts). Compounding their concern is the fact that most electricity markets are characterised by limited, oligopolistic competition in generation at best, and it is precisely when supply conditions are tight that such generators are able to exploit their market power (e.g. by strategically withholding capacity), thus exacerbating the heights to which wholesale prices rise. Their concerns are proven well-founded when deficiencies in market design add to such problems (such as gaming of the capacity price by generators in the now-abandoned British pool), or lead to outright failures (such as the Californian experience in 2000/01).

Additionally, real-world electricity prices should be expected to be only imperfect reflections of their theoretical counterparts. After all, all electricity markets are to some extent artificial, being borne of centrally-planned and controlled transitions, however complete, from regulated and otherwise centrally-planned and controlled electricity systems (e.g. vertically integrated utilities in the US). Furthermore they persistently attract strong political attention, whether from reform sponsors eager to avoid the adverse personal consequences of real or perceived failures, or from detractors eager to find fault. Importantly, all markets are subject to external regulation to some degree, well-intentioned and applied or otherwise, and this very fact can be sufficient to undermine their efficacy in balancing demand and supply, and eliciting investment.

Electricity price caps, for example, are often applied, at arbitrary levels or based on shaky, administrative estimates of the market value of unserved load (i.e. VOLL), as a means to constrain the exercise of generator market power during tight supply conditions.⁵ While such interventions may well reduce social costs from one form of market imperfection, market power, they do so at the expense of distorting important price signals, particularly when those signals are needed most. Thus the social costs of market power may be mitigated (or simply expressed otherwise), but consumers are also denied information about the economics of reducing demand or making investments in energy efficiency and flexibility. At the same time electricity markets risk not being able to match supply and demand (necessitating further, typically non-market based interventions to ration supply). Owners of existing generation assets are denied an important source of funding for their capacity investments, and new investors are denied the signals and profit opportunities provided by price spikes. Furthermore, generators are denied critical signals and incentives affecting their choices affecting availability rates of existing or new capacity (such as maintenance timing and levels, and technology choices). Not only do such price caps, if set too low, inherently prejudice the ability of electricity markets to supply demand at allowable prices, but by making generation investment less attractive they also reduce capacity margins and contribute to a greater probability of instability and forced rationing at any price. As Oren (2000, p. 6) puts it, "In some cases regulatory intervention in adequacy assurance is needed

⁵ Under the British pool VOLL was set administratively at £2,000/MWh in 1990 and then increased annually at the RPI (Roques et al. (2004)). In the Australian National Electricity Market an initial price cap of A\$5,000/MWh was subsequently doubled in 2002, and to enhance security of supply a further tripling was recommended in 2004 (*Triple electricity price cap: report*, The Age, theage.com.au).

to compensate for regulatory interference in the energy market . . . [arising] due to misperceptions and difficulties in distinguishing between market power and legitimate scarcity rents.” Indeed, according to Roques et al. (2004, p. 41), “The most convincing argument against energy only markets [i.e. that other mechanisms are required to encourage generation investment] lies in the difficulties for regulatory authorities to distinguish between the exercise of market power and legitimate scarcity rent.” Oren further acknowledges, however, that “Often the threat of regulatory interference to curb scarcity rents is sufficient to inhibit capital formation and raise the capital cost for investment in generation capacity.” (p. 6).

Such considerations contain three important lessons. The first is that actual electricity market performance in eliciting investment is not likely to be of itself sufficient to reassure reformers that liberalization is working as it should. This is not helped by a curious inability of many electricity market observers to appreciate how the market can and should work, despite similar markets in other industries providing useful analogues (although see the following section regarding electricity distinctives). The second is that many of the reasons why markets might not be trusted to deliver the desired outcomes are often based on artificial or discretionary grounds. If regulatory risk is hindering required investment, for example, should regulatory discretions be reduced (Helm (1994)), or further regulations be implemented to compensate for the adverse consequences of existing ones (e.g. price caps or capacity markets)? The third is that many of these concerns about electricity markets do not themselves imply that electricity markets cannot deliver the desired outcomes. Rather they suggest, with an evidentiary basis or not, that electricity markets do not do so. The questions then should be, is this due to inherent features of electricity markets, what can be done to compensate for any market deficiencies (inherent or not), and are the alternatives to electricity markets superior?

4.3. Explicit Concerns about Market Failure

The most serious critique of the view that energy-only electricity markets are adequate to elicit socially optimal electricity investment comes from an investigation of factors which may give rise to so-called “market failure.” Economic theory predicts that markets under ideal conditions will provide optimal outcomes, but also that when there are departures from these ideal conditions this need not remain the case. In the present context the question is then whether there are inherent features of electricity markets giving rise to market failure, and specifically, to underinvestment in generation capacity required to ensure the socially desirable level of security of supply/adequacy? To the extent such failures are identified, it is then necessary to consider what interventions, if any, will improve the situation (see Section 5).

The suggested sources of potential market failures are many. First are features of electricity systems and markets, some peculiar and some not, that are argued to mean the standard economic results cannot be assumed. Others relate to the market power often held by oligopolistic generators and the effect of market rules, regulation and industry structure. Often-times the suggested sources of market failure are in fact little more than government or regulatory failure being projected onto markets. Surprisingly, it has not been until relatively recently that some of the long-standing attitudes towards the nature of electricity provision inherited from the pre-liberalization “public service” model have been critiqued and dismantled. These themes are developed below.

It is commonly pointed out that electricity and its provision is different to that of other commodities. Unlike most commodities, it is currently uneconomic to store price-buffering inventories of electricity, so instead it must be produced and consumed in real-time. Also, unlike other commodities which can be physically identified wherever they are transported, the use of AC networks, with their associated physical laws, to transport electricity means that it is impossible to identify which given unit of electricity was produced by a specific generator and then consumed by a specific customer. Indeed, in the normal course it is

impossible to even control which customers get to consume electricity at any point in time, since short of switching off large groups of interconnected customers at once the only thing a consumer need to do to acquire electricity is plug in and turn on. This makes it impossible to ensure that actual flows of electricity correspond to those contracted for by producers and consumers. Moreover, the physics of interconnected electricity networks means that the action of any one connected party can simultaneously affect the quality and availability of energy to all other connected parties (somewhat like the systemic risk associated with wholesale payment systems in the bank sector). For this reason the provision of real-time ancillary services (e.g. voltage regulation, frequency support, spinning reserves) remains an important part of electricity market architecture (see Stoft (2002)).

There are features of electricity markets in addition to these inherent characteristics of electricity provision that are also argued to be important distinctives. Foremost is the relative lack of price-responsiveness in electricity demand (electricity demand has a low or zero price elasticity), particularly in real-time but also over longer time-frames, not to mention its high level of short-term unpredictability. In part this is because current electricity networks commonly inherit metering technology allowing only ex post charging, for all but the largest customers, for consumption over long time periods (e.g. monthly). It also arises when wholesale electricity market arrangements do not allow for demand-side bidding, instead relying (as in pre-liberalized systems) on system operator demand forecasts. However, even where customers are price-responsive they often hedge their price exposure via financial contracts, thus dampening their exposure to price signals. Indeed, Roques et al. (2004) report that while around half of British electricity demand is on real-time meters able to measure and charge for consumption on a continuous basis at prevailing energy prices, most customers hedge their price risk anyway by buying on contract.

This lack of demand price-responsiveness heightens the risk that demand will not be able to be equated with available supply under adverse conditions at any price, threatening system stability and raising the risk of administrative rationing, and suggesting that greater capacity is therefore required to ensure adequacy. This is worsened by the inability, given current network configurations, to selectively curtail most, smaller customers' demand even if they should be prepared to allow this when needed.⁶ The importance for investment of demand unresponsiveness to price may be overstated, however, with simulations by Hirst and Hadley (1999) showing that only a small fraction of load needs to respond to real-time prices in order for energy-only market driven investment to be sufficient to maintain system adequacy. While it is commonly argued that price inelasticity of demand is a critical flaw in electricity markets (Hobbs et al. (2001), Bidwell and Henney (2004)), the importance of this aspect remains an empirical question so long as most large customers, or at least some sizeable customers, are price-responsive (as many are, contracting for interruptible load as part of ancillary service provisions and otherwise).

At the same time short-term electricity supply is similarly price-unresponsive, particularly when available generating capacity is nearing its immediate limit. While the price elasticity of supply may be positive when supply is well-within available capacity, or over the longer-term when plant can be brought on-stream and/or created, there is only so much generating capacity available at any instant in time to service demand irrespective of market prices.⁷

Added to this is the fact that electricity generation technologies are such that it is currently uneconomic for more than a few major generators to operate in any given market, and limitations in transmission capacity make it implausible or uneconomic to import or export sufficient energy over long distances to enable unfettered competition in generation (within or across markets). Hence generation in any one electricity market is typically characterised by

⁶ Although most small electricity customers in New Zealand, for example, have ripple control installed in electric water heating, allowing direct load control.

⁷ The same can also be said, however, of fresh produce once harvested and delivered to market for sale (e.g. lettuces). The supply quantity on that day is then fixed, and limited storability means inventories cannot significantly smooth production and price swings over time. Sudden shortages in fresh produce, such as due to unseasonably poor weather, can thus lead to significant price swings to balance supply and demand.

only a few competitors enjoying some degree of market power. The usual economic prediction in this case applies – supply levels will be lower, and prices higher, than the welfare-maximising levels predicted for the ideal (but notional) benchmark of perfect competition.

Furthermore, outside of real-time spot or day-ahead electricity markets, there is little depth or “liquidity” in markets for forward electricity trading. It is commonly held that this illiquidity is an obstacle to long-term generation investments (de Vries and Hakvoort (2004)). Certainly the security of a long-term contract would make generation investment more viable, but the lack of such contracts does not imply that it is not. As stated by Roques et al. (2004, p. 33), “looking at other commodity industries . . . one notices that investments are made on the basis of expectations and not thanks to binding forward contracts.” That being said, despite the lack of liquid, long-term forward markets for electricity it is not unusual for generation projects to be built (e.g. project financed) on the strength of non-traded, long-term supply contracts with major electricity users in need of enduring, price-hedged supply. Projects such as Loy Yang B in Victoria, Australia, with long-term supply contracts to aluminium smelters, being a case in point. Contractual or structural arrangements that provide long-term certainty to generation investment returns should, all other things being equal, usefully complement the role of spot electricity markets in eliciting investment. However, a lack of liquidity in traded, long-term forward contract markets does not mean long-term contracts are not being used to elicit generation investment.

It is also suggested (for example, de Vries and Hakvoort (2004)) that difficulties in forecasting key electricity market variables, in part due to the absence of long time series on these variables, is a further obstacle to electricity investment. Once again, however, this could be said of other industries, such as telecommunications (for example, regarding cellular phone or broadband internet take-up). An inability to forecast such variables has not proved an obstacle to significant investments in these areas. Certainly forecasting difficulties cannot assist generators’ long-term investment decisions, however developments in financial decision-making over the past 20 years, specifically in the application of real options techniques in investment evaluation, provide a means to accommodate key uncertainties (see, for example, Small (1999), Botterud and Korpas (2004)). Indeed, the inclination to rely on forecasts and detailed industry modeling for investment planning bears more than a little resemblance to the centralized coordination and control approach pre-liberalization (though, to be fair, sophisticated market participants may wish to use such methods too, albeit with strategic rather than central-planner optimizations).

Finally, liberalized electricity industries have typically involved the enforced separation of previously integrated activities. This is generally to facilitate non-discriminatory grid access by competing new generation, remove anti-competitive cross-subsidies from monopoly distribution and competitive activities, and to thereby support and encourage competition and private investment. The development of wholesale electricity markets has been a key structural innovation in addition to these structural separations. Relevant here are the informational asymmetries and scale and scope diseconomies to which such structural separations give rise. With integrated electricity providers it is possible to coordinate activities and investments at a technical level, use ownership as a means to mitigate (or exploit) upstream or downstream market power (Coase (1937)), and to exploit cost efficiencies achievable with scale and where there are overlaps in related activities (Teece (1980)). Thus it is possible that changes to industry structure implemented as part of electricity liberalization may give rise to problems that integration had solved (by design or good fortune) and which have the potential to hamper market performance. Work by Benz et al. (2003) suggests that while integration can be beneficial for investment incentives, this is not unambiguously so.

These features have some important consequences. With both demand and supply being relatively price-unresponsive, particularly when demand is high relative to demand (and because electricity cannot be economically stored), small changes in either supply or demand can require dramatic changes in spot electricity prices to ensure real-time balance is maintained. This implies a greater need to use financial hedge instruments (or structural

measures) to manage price risk than is the case for other commodities (for which inventories can be used to buffer price swings). The resulting volatility in spot prices can also be exacerbated by the exercise of generator market power (e.g. strategic withdrawal or withholding of capacity to secure higher prices), but also by actions of the system operator to maintain stability during tight supply conditions - "under certain contingencies the market price, and the associated scarcity rents available to support investments in generating capacity, are extremely sensitive to small mistakes or discretionary actions by the system operator." (Joskow and Tirole (2004), p. 47).

Thus it is difficult to discern which part of high electricity prices is due to genuine scarcity (providing the all-important scarcity rents required to fund generation investment), which part is due to mistakes and discretions of the system operator, and which is a reflection of market power (with price being above and supply being lower than their socially optimal levels)? While excessive peak pricing should be expected to encourage investment in capacity, thereby dampening future price volatility, interventions to address market power may be quite critical – e.g. since investment in peaking plant relies on peak prices for its viability, "small distortions of the investment signal may have large consequences" (de Vries and Hakvoort (2004, p. 4)). However, the perception of at least the possibility that electricity price spikes are due to generator market power raises the significant risk that regulators will face political pressure to intervene, with the likely consequence that critical investment signals and incentives will be distorted – "theory and experience indicate that reliance on energy prices alone to motivate sufficient capacity is economically and politically risky" (Hobbs et al. (2001, p. 32)).

More significantly for investment and adequacy, authors such as Hobbs et al. (2001) and Bidwell and Henney (2004) argue that the lingering lack of price-responsiveness on the demand side means that electricity markets are failing to accurately signal consumer preferences for adequacy (their "willingness to pay"). Hobbs et al. argue that extreme electricity price spikes are more a reflection of electricity retailers' unwillingness to suffer political criticism in response to forced outages than they are of consumers' willingness to pay for ongoing supply. This is a strong judgement, especially since they also acknowledge consumer resistance to the adoption of real-time pricing (consistent with British experience reported by Roques et al. (2004), and Doorman's (2003) concern that spot customers would probably hedge their price exposure anyway). While it is true that greater demand-side price-responsiveness would reduce the risk of system instability and longer-term supply inadequacy, as well as reduce spot price volatility and exposure to manipulation by generators with market power, the premium that consumers are prepared to pay in their fixed price contracts over average spot prices should provide a measure of their willingness to pay for price security.

Conversely, examples abound of mechanisms that successfully encourage consumers to reduce consumption during shortages even without real-time price exposure. Where supply shortages are predictable and/or long-lived, calls for voluntary public savings campaigns such as those in California in 2000/01 and New Zealand in 1992, 2001 and 2003 have elicited significant energy savings. Despite some official resistance in New Zealand to future such campaigns (Ministry of Economic Development (2004)), they are likely to be cheaper than installing new hardware or building new generation. In some cases customers on monthly billing have been offered rebates if their monthly demand during the crisis is below some measure of average demand for that period to further encourage energy savings. In the more normal course, many residential customers in New Zealand are able to opt for pricing plans that offer a discount to those willing to have electric hot water heating curtailed by suppliers via ripple control. Additionally, as energy exchanges develop for successively smaller customer classes, the offering of fixed volume supply contracts with top-ups or sell-backs via such exchanges (or at supply contract imbalance prices) could be expected to encourage greater price-responsiveness across all customer classes. Hence concerns about the inability of market arrangements to either elicit demand reductions when necessary or consumers' willingness to pay for supply security are possibly overstated.

Based on the discussion thus far, few of these features of electricity systems and markets are likely to be inherent causes of underinvestment in generation leading to socially suboptimal security of supply. Furthermore, market power in generation might be regarded as a necessary evil given current technologies (with the inherent effect of this evil on investment levels being ambiguous, and possibly beneficial). So too might the political and regulatory inclination to intervene in electricity markets in response to market power or other concerns (such as price volatility, which too is not inherently bad for investment). However, in this case the risk to investment is discretionary and institutional rather than technological, and hence suboptimal investment due to interventions are more a sign of government failure than market failure (see, for example, Friedman (2004), Laffont and Tirole (1991), Joskow and Rose (1989), Stigler (1971), Helm (2004)). So to fully fathom the reasons for possible failure in electricity markets we must look further still.

The most promising criticisms of electricity markets, suggesting they will under-provide capacity and security of supply, involve “public good” and “externality” arguments (examples of so-called “market failure”). A “private good” is both “excludable” (i.e. free-riding can be precluded, with only those willing to pay for a good receiving it) and “rivalrous” (one person’s consumption of a good or service reduces the amount available for consumption by others). Under such circumstances it is possible for profit-motivated competing suppliers to vie to supply that good or service and secure payment for doing so, and at the same time consumers indicate their willingness to pay for such goods or services by buying them. A “public good” is usually characterised as one which lacks these characteristics, and is often also non-rejectable (i.e. if it is provided, consumers must consume it).⁸ Thus a public good or service is non-excludable (the classic examples cited are unscrambled public broadcasts and lighthouses – once you have them it is hard to stop people from benefiting from their services) and non-rivalrous (one ship seeing the lighthouse won’t stop another from seeing the lighthouse too). National defence is an example of a non-rejectable public good. In these cases consumers have an incentive to free-ride on the production of such goods or services, and do not reveal their true willingness to pay for them. Markets may then be unable to provide those goods or services to the socially optimal level, with their provision instead often falling to the state, funded through compulsory levy (i.e. taxation). Thus, if left to market provision, the risk is that the socially desirable level of production of those goods and services will be undersupplied (an outcome analogous to that of “prisoners’ dilemma” – Friedman (2004)).

Security of supply/adequacy (but not energy provision) is argued by various authors to have the characteristics of public goods (e.g. de Vries and Hakvoort (2004)).⁹ Due to network physics, any party investing in new generation capacity (or otherwise contributing to a lower risk of network collapse or outage) cannot capture that benefit itself, since any other connected party cannot be excluded from consuming the extra supply. Such a view is evident in the US Federal Energy Regulatory Commission’s resource adequacy specifications in its proposed standard market design (Rochlin (2004)). However, adequacy clearly does not share the non-rivalrous characteristic of public goods (Bidwell and Henney (2004), Rochlin (2004)). The difficulty of electricity networks is that the consumption decision of one party affects the supply security of others – the more I consume the greater the risk that you will not be able to. Furthermore, Rochlin (2004) points out that non-excludability itself is not the source of potential market under-provision of adequacy. Citing the famous critique offered by Coase (1974) of the lighthouse example of public goods, he instead

⁸ *Penguin dictionary of economics*, 7 ed.

⁹ Indeed, de Vries and Hakvoort (2004) describe five sources of what they call “market failure” that contribute to adequacy underprovision where only spot electricity prices are used to signal and fund required investments by private parties. Specifically they cite price restrictions, difficulties in eliciting consumers’ willingness to pay for security, regulatory uncertainty, regulatory investment restrictions and investor risk-aversion. The willingness to pay problem has already been argued to be overstated. Investor risk-aversion is not a source of market failure, but a characteristic of investors (and there are difficulties in arguing that taxpayers are less risk-averse regarding public investments than private investors are regarding private ones). The other three problems all involve regulatory impositions, and thus are more properly characterised as government or regulatory failures, rather than market failures.

identifies the inability of parties to secure payment for the relevant good that affects its private delivery.

So long as property rights are specified or some other mechanism implemented to make consumption chargeable (in the case of lighthouses by enabling owners to charge port fees for landing vessels, or in the case of public broadcasts through scrambling technologies that require excludable decoders), it is still possible for private parties to profitably supply the non-excludable goods or services without resort to state intervention and tax-based funding. Indeed, such measures can be sufficient but are not necessary to ensure this is so, since, for example, even free-to-air radio can be profitable to private providers so long as advertisers are prepared to pay for broadcast advertisements (Friedman (2004)).¹⁰ The challenge in the present context is to make these measures chargeable in a way that is not simply imposed on consumers, but rather enables them to opt for their desired level of security for a commensurate price (e.g. see the discussion of load-limiting fuses in Section 5).¹¹

Thus the public good argument for private undersupply of adequacy, and rationale for corrective interventions, is increasingly regarded as weaker than first thought, and perhaps merely a reflection of the earlier “public service” model of electricity supply.¹² Indeed, Oren (2000) argues that while real-time security has public good characteristics, longer-term adequacy is simply a private good, being private insurance against shortages. The case for intervention then requires further grounds for market failure. Various authors suggest that this ground lies in the “externalities” associated with adequacy provision (e.g. Jaffe and Felder (1996), Bidwell and Henney (2004), de Vries and Hakvoort (2004), Rochlin (2004)), once again reflecting electricity network physics.

Externalities arise where private costs or benefits from an activity diverge from their social costs or benefits, meaning that decisions made by private parties will diverge from their social optimum because such parties typically take only their private costs or benefits into account. Unlike for public goods, consumers’ willingness to pay can still be elicited when there are externalities. In the case of security of supply/adequacy, there is a positive externality (meaning social benefits exceed private benefits) from investments and other actions that increase adequacy – e.g. while I might secure sufficient private benefits to invest in new generation or energy-efficient technologies, the social benefit to me doing so may exceed my private benefits due to the enhanced adequacy that accrues to all users.¹³ Thus I would under-invest in those things from society’s standpoint. Alternatively there are negative externalities argued to be associated with electricity consumption decisions – if I consume a certain amount of electricity given the direct cost to me of that electricity, I fail to account for the extra social cost (i.e. to other consumers) arising from the increased risk of supply instability or inadequacy to which it gives rise.¹⁴ This mixture of production and consumption

¹⁰ Various authors point out that even genuine public goods do not necessitate state provision. Just as governments often supply private goods, there are numerous instances where private parties provide public goods (e.g. Friedman (2004)).

¹¹ Ostrom (2000) defines resources which are rivalrous but non-excludable (implicitly assuming they are also not chargeable) not to be public goods, but instead “common pool resources.” She identifies factors contributing to the success of endogenously-determined management regimes for common pool resources without the need for regulatory intervention (indeed, the superiority of endogenous over imposed regimes). If charging mechanisms are not implemented to support the private provision of electricity security of supply/adequacy (should support be considered necessary), adequacy can be treated as a common pool resource and such management regimes would then merit investigation.

¹² The appetite of some economists for finding public goods and arguing for market failure is such that electricity is not alone in this regard. Even oil reserves have been argued to be a public good, requiring imposed adequacy levels – despite their clear excludability and rivalrousness, not to mention storability (Colegrave et al. (2004)).

¹³ A scenario not often considered is that where new generation produces electricity so cheaply that its commissioning suppresses electricity prices and displaces a greater capacity of more expensive existing capacity, diminishing adequacy. Also not considered in externality arguments relating to adequacy are scenarios where new generation investments increase grid congestion, a negative externality. The use of financial transmission rights to internalise such externalities is discussed in Evans and Meade (2001).

¹⁴ Of course, to the extent my consumption decision increases electricity prices, which is more likely with price-inelastic supply and demand, this helps to elicit new generation investment that should enhance

externalities suggests that privately-motivated investment in adequacy (on both the supply and demand sides) funded only by energy-only electricity prices will fall short of the socially optimal level. While conceptually this argument appears sound, it is an empirical question whether these externalities are material in practice, of sufficient scale to require mitigation, or worse than any measures adopted to remedy them. In principle the superiority of interventions to remedy these externalities can be questioned (see Section 5), but the author is not aware of any evidence regarding these empirical questions.

The usual approach to mitigating externalities (assuming they are important) is to have them “internalised.” Traditionally this involves the use of taxes or subsidies to ensure consumers or producers bear a greater share of the true costs and benefits of their actions. Increasingly, more sophisticated and market-based tools are used, such as tradable emission rights to internalise to emitters the social costs of their pollution. In any case, Rochlin (2004) suggests that energy-only electricity markets are in fact sufficient to internalise any adequacy-related externalities. Since electricity prices spike during times of shortage, those consumers facing those prices face the immediate cost of supply and choose whether or not they wish to bear it. (Those on fixed-price contracts do not face that signal in real time, but over time face average fixed prices reflecting occasional spikes and hence face a derived price signal encouraging efficiency and conservation.) Where market prices do not ration available supply and unselective outages are forced (which begs the question why not?), he suggests it is possible to internalise the externality forced on customers of retailers with sufficient supplies interrupted due to the actions of retailers with inadequate supplies. In this case penalties could be imposed on retailers who are net purchasers during shortages, with the proceeds distributed to the customers of retailers who are net suppliers. While a simple intervention of this nature might work, problems remain in defining optimal penalty levels and customer shares of those penalties, not to mention in assessing the associated risk of supplier bankruptcy (which would already be high for net purchasers during price spikes – see Section 6.4). Attention here should instead be focused on remedying the failure of prices to balance supply and demand in the first instance.

Finally, work has been done to examine the dynamics of electricity investment under the energy-market only approach. Ford (1999), for example, models “boom-bust” investment cycles under market rules like those existing in California before the 2000/01 melt-down. Consistent with predictions of other authors (e.g. Bidwell and Henney (2004)) this suggests that when generation investment is left to private parties, signalled and funded by only spot electricity prices, cycles of tightening supply, declining adequacy and increasing prices (and volatility) will lead to sporadic investment after which adequacy and prices (and volatility) will sharply correct, only to subsequently recur. Mechanisms such as capacity markets are suggested as a means to dampen and smooth such cycles, with benefits predicted to include greater political acceptability of the resulting prices and adequacy. Less work has been done to assess the similarly smoothing impact of hedging via long-term financial contracts (where they are available) or structural measures such as vertical integration. In any event, no evidence is offered that such cycles are unusual where large, irreversible and long-lived investments are involved, or inherently welfare-reducing.

5. INTERVENTIONS INTENDED TO CORRECT FOR UNDERINVESTMENT

5.1. Proposed and Implemented Mechanisms in Brief

A full review, description and appraisal of the various mechanisms and interventions suggested to elicit socially optimal generation investments and ensure adequacy are beyond the scope of this work. Examples of suggested mechanisms can be found in Oren (2000), Hobbs et al. (2001), Vazquez et al. (2002), Creti and Fabra (2003), Doorman (2003), Fraser

adequacy. The point here is that externalities are utterly commonplace. The interesting question is their importance in practice.

and Lo Passo (2003) and Joskow and Tirole (2004). A brief comparison of approaches is provided by de Vries and Hakvoort (2004). This appraisal considers the essential features of such mechanisms as they affect their merits relative to reliance on energy-only electricity markets to elicit socially optimal adequacy. However, a brief description of each is provided.

Following Creti and Fabra (2003), common capacity mechanisms can be broadly divided into two classes – price-based and quantity-based. Price-based mechanisms are variously described as “capacity payments” (Joskow and Tirole (2004), de Vries and Hakvoort (2004), Oren (2000)) or “operating reserves” (Hobbs et al. (2001)). Under this approach generators are rewarded for capacity availability via lump-sum payments (Argentina, Spain), or via an uplift on energy prices (England and Wales pool pre-NETA), with payments being initiated by the market or system operator when operating reserves fall below some target level. While the capacity price/subsidy is fixed, the amount of generating investment elicited is left to market participants to determine.

Quantity-based mechanisms are variously known as “planning reserves” (Oren (2000)), “operating reserves” (Creti and Fabra (2003)), “installed capacity markets” or ICAPs (Hobbs et al. (2001), Creti and Fabra (2003)) and “capacity requirements/obligations” (de Vries and Hakvoort (2004), Joskow and Tirole (2004)). Under this approach load serving entities (i.e. energy retailers) are required to retain (or contract with generators to provide) a system-operator prescribed capacity margin over peak load, with the associated costs passed on to their customers. Thus the capacity requirement is fixed, with the cost of this then being market-based.

Other mechanisms include options-based schemes (Oren (2000), Vazquez et al. (2002)), capacity subscriptions (Doorman (2003), Roques et al. (2004)), retention of consumer franchises (de Vries and Hakvoort (2004)) and regulating for long-term hedge contracts (New Zealand’s Electricity and Gas Industries Bill, enacted in December 2004). Under options-based schemes the system operator purchases options from generators to secure a particular level of system adequacy, requiring generators to pay the operator the market price less the strike price when called. Generators able to operate when the options are called thus receive the strike price (i.e. are hedged against price movements for the contracted capacity), whereas those unable to provide the contracted capacity bear an ex ante uncertain cost. This encourages the generators to match capacity with options written, and also to ensure availability during periods of shortage when the system operator is likely to exercise them.

Doorman’s capacity subscription scheme is probably the most market-oriented of the alternatives currently debated, coming at the cost of requiring new system hardware rather than new market arrangements. Instead of attempting to induce greater price-responsive demand through more widespread installation of real-time metering (which, as discussed, is neither necessary nor sufficient to do so), this scheme requires the installation of load-limiting devices (LLDs, in this case fuses rated at a particular capacity) that can be triggered by the system operator during shortages. Customers would buy fuses for their preferred level of load, thus subscribing to a particular level of ongoing supply at a price.¹⁵ By introducing the ability for demand to be selectively curtailed (unique to a given level of fuse capacity), and allowing customers to choose their desired level of supply security, this mechanism clearly elicits their willingness to pay for supply security and eliminates free-riding. Total fuse sales signal required reserve capacity requirements, and also provide revenues to support investment. Security of supply under this approach is thus a decidedly private good.

Retaining consumer franchises, by contrast, meaning that energy retailing companies continue to enjoy captive customer bases and can therefore pass on capacity charges to their consumers, is a simple but blunt means to fund investments deemed necessary for adequacy. Such an approach sits uncomfortably with electricity liberalization, in that it denies customers choice of energy supplier, imposes adequacy levels, and creates market power instead of competition. Past experience indicates it need not guarantee adequacy, or if it

¹⁵ This is a more sophisticated version of the hot water heating ripple control mechanism discussed earlier and common in New Zealand.

does, at the risk of overcharging. Legislating that generators must offer a fixed proportion of their output via long-term contracts, and that wholesale purchasers of electricity must similarly do so with some proportion under long-term contracts, is an option being considered in New Zealand and allowed for under recently-enacted legislation. It is doubtful that such moves will make generation investment any more attractive under New Zealand's current industry structure, in which 91% of generation is vertically integrated with and thus naturally hedged by energy retailing (Electricity Commission (2004)), but would rather force a certain level of de-integration and create a tenuous and artificial rationale for stand-alone energy retailing where it is not otherwise viable (Evans and Meade (2005)).

5.2. Capacity Mechanisms Compared and Critiqued

The usual caveats regarding the use of models in drawing policy conclusions are obviously in order. Not only are they simplifications of inherently complex phenomena, but they rely on assumptions and parameter estimates that often differ from reality. In the present case it is important to acknowledge the importance of regulatory risk, difficulties in the estimation of key parameters (such as VOLL), ubiquity of market power due to oligopolistic generation, and demand (and supply) price elasticity which, if low, is not generally zero. Predictions will be less reliable for models that do not allow for these than for models that do.

Hobbs et al. (2001) simulate generation investment incentives under energy only markets, ICAPs and operating reserve markets, assuming a common level of adequacy in each case. They find that, under conditions that are both ideal (e.g. no market power) and extreme (completely price-unresponsive demand), price- and quantity-based capacity mechanisms can – like energy-only markets – induce the required level of generation to assure the target adequacy level. They do so, however, with lower price volatility and smoother generator profits (reducing investment risk), and lower price caps (thereby constraining any market power), than energy only markets. These predicted benefits form the main normative basis for suggesting capacity mechanisms. This result is achievable, they argue, without inflating generator profits or distorting market prices, although they caution that this is not the case if capacity mechanisms distort maintenance incentives (causing divergences between capacity and availability), or if demand faces real-time electricity prices.¹⁶ Moreover, they acknowledge that actual capacity markets suffer various deficiencies (such as the ability of PJM generators to export previously-committed capacity to other markets, or gaming of ICAP auctions in New England), some of which can be remedied by market refinements. Other problems are more enduring, they admit, such as the self-defeating feature of price caps which often accompany capacity markets, given they muzzle price signals during shortages that otherwise would simultaneously encourage both energy efficiency/conservation and new generation investment. Finally, they acknowledge that critics regard capacity markets as being “an unnecessary relic of regulation that provides nothing of value . . . defining artificial markets that give generators additional opportunities to manipulate prices and extract underserved revenues” (p. 24). They conclude, however, that capacity markets are likely to be more politically acceptable than energy-only markets, and may improve public confidence in electricity restructuring post-California (adding further to their rationale).

In contrast to these pragmatic arguments and conclusions, Creti and Fabra (2003) develop a capacity model examining the welfare implications of their introduction, as well as their optimal configuration assuming they are introduced. They find that capacity markets do not necessarily maximise social welfare, as the costs of full reliability may exceed the gains from avoiding rationing, particularly when the opportunity costs of generators' commitments

¹⁶ It is notable that they find that ICAP markets do not subsidise capacity that is otherwise uneconomic. A reserve generation scheme was implemented in New Zealand in 2004 funded by a compulsory uplift that involved the re-commissioning of plant by government that had not long previously been decommissioned by its owner.

are large (as is the case in PJM, for example, when it can be more profitable to export energy than relieve local shortages). Whether or not it is optimal to introduce a capacity market can hinge on the degree of competition in generation, with introduction possibly optimal under monopoly, but suboptimal with sufficient competition. If a capacity market is introduced, departures from indicated optimal parameters (such as price cap or capacity levels) undermine its function and/or leaves scope for market power exploitation by generators. Indeed, they note reports from PJM that market power is structurally endemic to its capacity markets, which mirrors gaming problems reported in the abandoned England and Wales pool (i.e. strategic withholding of capacity to increase LOLP and hence capacity payments).

Joskow and Tirole (2004) similarly model optimal configurations of capacity markets and accompanying price caps. Noting that rationing may in some circumstances be socially optimal, they point out that the rationale for capacity mechanisms must therefore be other than just to avoid rationing altogether. Oligopolistic generation is identified as a potential source of electricity price distortion, and an optimal price cap level is derived. However, if actual price caps are set below their optimum it becomes necessary to make capacity payments to recover lost incentives for investment in peaking plant. A complication they identify is that this then raises the possibility that generators with market power simply exploit that power *ex ante* via the capacity mechanism and *ex post* price cap.¹⁷ The same investment and adequacy levels as deriving under no price cap and with energy-only markets can then result. Moreover, Joskow and Tirole find it is possible that price caps combined with capacity payments are insufficient to restore optimal pricing in some circumstances. Finally, they show that when a system operator contracts for peaking capacity funded by an energy price-uplift, beyond a point this undermines the profitability of private peaking plant and thus crowds out private investment, leaving the only peaking investments to be undertaken by the system operator. This is a potential risk in New Zealand, with reserve generation having been contracted for by the regulator at a relatively low trigger price (i.e. *de facto* price cap) of NZ\$200/MWh (i.e. US\$140/MWh at January 2005 exchange rates), with no other generation capacity incentives in place.

At a more operational level, both price- and quantity-based mechanisms typically involve the estimation of VOLL and LOLP. In theory both types of mechanism should be equivalent (Jaffe and Felder (1996)), but this equivalence fails when LOLP and VOLL are mis-estimated (or simply not used to determine optimal capacity margins). Indeed, LOLP and VOLL commonly rely on engineering-based estimates, without reference to market-based parameters (except perhaps where auctions are held by system operators to secure interruptible load), or are even set arbitrarily. Roques et al. (2004) cite research from the abandoned England and Wales pool that LOLP was almost certainly overstated, while VOLL was probably underestimated (with errors somewhat cancelling). This implies that any capacity secured by these mechanisms cannot be assumed to represent the socially optimal level, particularly since the risks of over-estimation can be argued to exceed those of under-estimation (Jaffe and Felder (1996)), and because regulators and politicians have an incentive to over-specify required margins, as this reduces their political downside from outages and because consumers, not they, bear the costs of overcapacity (Rochlin (2004)). Jaffe and Felder suggest that policy focus should be directed towards correctly estimating the socially optimal level of reserve generation, noting that there can be no presumption that an artificially determined level is optimal, and that setting the wrong reserve margin can easily be worse than simply ignoring the externality. Indeed, they state that “to be sure” the

¹⁷ A curious analogue to this idea can be found in a pricing debate held in New Zealand in 1991 (Electricity Corporation of New Zealand (1991)). The then monopolist generator attempted to raise prices towards long-run marginal cost (including amortised capital costs) on the argument that this was necessary to fund new investment (akin to return-regulated generators seeking an uplift on marginal cost to cover capital costs of investment, absent wholesale markets and price spikes providing scarcity rents). This failed move by the monopoly generator might be construed as an attempt to *ex ante* capitalise scarcity rents, rents to market power, or both, into then administratively determined wholesale electricity prices.

then-existing capacity mechanism examples did not equate marginal social benefit and cost when determining reserve margin levels.

An important consequence of misestimating optimal capacity levels when implementing capacity mechanisms is that this then distorts both capacity and spot (and forward) energy prices. For example, if capacity prices are set too high, this encourages overinvestment in capacity, depressing energy prices and encouraging over-consumption of energy (which in turn fuels greater need for capacity). As Oren (2000, p. 2) puts it, there is a “strong possibility that the measures taken to ensure generation adequacy have the effect of suppressing energy prices due to excess capacity or perverse incentives so that the necessity of such measures becomes self-perpetuating.” He cites the Argentinian experience, with capacity payments that induced generators to bid below marginal cost to boost output and increase capacity payments, as an illustration. Moreover, the various approaches differently reward actual availability during peak demand periods, with ICAPs, for example, rewarding capacity (“iron in the ground”) rather than actual availability (Hobbs et al. (2001)). Energy-only markets and (despite their other shortcomings) capacity payment mechanisms, by contrast, better reward actual availability, and thus can also be expected to encourage more desirable maintenance policies.

Both price- and quantity-based mechanisms overcome public good and externality aspects of security of supply/adequacy. Capacity payments, for example, resemble the orthodox formula for internalising an externality, being little more than a tax on consumption and tax on supply. Capacity obligations similarly require that charges be imposed on consumers without eliciting their willingness to pay for adequacy. Like the options-based mechanisms, they continue to require centralized determinations by the system operator or regulator as to desirable levels of adequacy and capacity. Retaining customer franchises and forcing levels of long-term contracting do so even more bluntly. Capacity subscriptions, with their relatively inexpensive hardware requirements, to date offer the most elegant solution. They simultaneously ensure adequacy is a private good while eliciting consumers’ willingness to pay for adequacy. To the extent that energy-only electricity markets are unable to elicit socially optimal adequacy, the capacity subscriptions approach, as well as other market-based mechanisms for securing interruptible load, would appear the least distorting methods currently available.

5.3. The Case for Intervention

Aside from the practical difficulties in properly implementing capacity mechanisms, it is important to consider their introduction more broadly. First it should be asked whether any residual public good or externality issues thought to cause inadequacy are in practice material. Part of this might involve investigating whether actual generation investment levels (and other measures affecting adequacy) are proving to be effective, and the International Energy Agency (2002) evidence cited earlier provides some assurance in this regard. Assuming doubts remain about adequacy under energy-only markets, it then needs to be asked whether available alternatives are superior, weighing both their expected benefits and costs. As Helm (1994, p. 18) puts it, “Intervention by government is only efficient if the costs of the market failures which it addresses exceed the costs of intervention. Market failures must exceed regulatory failures.” Indeed, it is important to ensure that any causes of “market failure” are properly attributed – “The kind of situation which economists are prone to consider as requiring corrective governmental action is, in fact, often the result of governmental action” (Coase (1988, p. 133)). The counter-productive effects of price caps are an important case in point.¹⁸ Some authors express the view that interventions to correct market failures are more at risk of failure themselves – “The circumstances leading to market failure are exceptional in the private market but normal in the public market, thus the

¹⁸ Indeed, Bidwell and Henney (2004) and Creti and Fabra (2003) also note the tendency for price caps to become self-fulfilling in the presence of market power, a potential regulatory “own-goal.”

existence of market failure is an argument for market solutions and against political ones (Friedman (2004, p. 18)). Indeed, if imposed energy-only electricity markets have not been successful, what reason is there to believe that imposed capacity markets or other capacity mechanisms will be any more successful and immune to the risks of policy failure or further policy creep? If the risks of intervention in energy markets are sufficient to hinder generation investment, why not in capacity mechanisms also?

It is equally important to ensure that problems identified in one market are not simply duplicated or merely shifted into another. The discussion above already identifies market power as one such problem that can be shifted or re-expressed with the introduction of capacity mechanisms. Less frequently identified is the problem of shifting regulatory risk. Authors such as Oren (2000), Hobbs et al. (2001) and Bidwell and Henney (2004) identify the investment disincentives posed by the threat of regulatory intervention to constrain high and volatile electricity prices. On the other hand, Fraser and Lo Passo (2004, p. 57) note in relation to the Spanish capacity incentive that “Uncertainty about future levels means that the incentive is not trusted by investors, although it has provided sufficient incentives to keep old existing plant open.” Similarly, just as spiking energy prices subject to allegations of market power abuse may be regarded as politically untenable, so too should be capacity payments similarly subject to market power abuse. This is particularly so if the mechanisms come to be viewed as providing safe generator profits at consumers’ expense, not least if they fail to deliver adequacy (which is inevitable at least some time), but also if they do (consumers ask why they need to pay to insure against something that does not appear to be a significant risk, Vazquez et al. (2002)).

Where politicians or regulators make themselves responsible for security of supply this will inevitably crowd at least some, and according to Joskow and Tirole (2004) possibly all, private investment in peaking capacity (if not base load capacity). Roques et al (2004) identify the precedent effect of government interventions, such as those in Britain to support loss-making generators. Such interventions induce moral hazard issues, with market participants coming to believe they need not implement their own adequacy measures in the expectation that government will ensure adequacy at all market participants’ expense. De Vries and Hakvoort (2004) correctly identify that regulatory risks represent a negative externality from policy changes (and indeed, interventions) that undermine investment. Even assuming the public good case for arguing market failure, Ostrom (2000, p. 138) notes that “Solid empirical evidence is mounting that governmental policy can frustrate, rather than facilitate, the private provision of public goods.” Indeed, she states that externally imposed solutions (e.g. re adequacy) “tend to crowd out endogenous cooperative behaviour” (p. 147), suggesting that market-based solutions to any underinvestment problem are then less likely to endogenously evolve – the visible (yet unpredictable) hand of regulation impeding the market’s “invisible hand.”

The latter point is fundamental, and discussed further in Section 6 regarding vertical integration as a possible, natural solution to investment concerns. Despite believing that energy-only electricity markets elicit inadequate investment, Hobbs et al. (2002) acknowledge that capacity markets are unusual, in that “Dairy farmers don’t peddle ‘udder capacity’ separately from milk, and consumers are not required to buy refinery capacity to gain the right to purchase gasoline” (p. 23). If capacity markets evolve of their own right in response to market incentives it would seem more likely that they are an economically efficient solution to a genuine problem. This evolution could be supported by consumers being able to seek compensation from suppliers in the event of outages, perhaps with tiers of guarantee available for different levels of insurance premium (cf insurance arrangements suggested by Roques et al. (2004)).¹⁹ This then motivates suppliers, who are better able than consumers to manage their adequacy risks, to institute arrangements to ensure they are not left liable for other parties’ actions. Where capacity mechanisms do not endogenously

¹⁹ It has taken 20 years of liberalisation in New Zealand for electricity consumers to finally enjoy the same consumer protections as they do at law for other goods or services. Supplier liability to most customers for supply interruptions has to date been limited or non-existent.

evolve, but are instead imposed by fallible regulation, their efficacy should naturally be in question.

Recent New Zealand experience illustrates how such crowding out can easily occur. A reserve generation scheme has been implemented with a very low trigger price and no separate investment incentive to compensate for its effective capping of peak spot prices. Having made itself responsible for supply security, government has warned industry that it will invest in new capacity if industry fails to deliver. The responsible government minister backs this threat by publicly stating “This is an activist Government and I am an activist minister.”²⁰ The new electricity industry regulator, the Electricity Commission, has had its discretions considerably widened under the Electricity and Gas Industries Bill enacted in December 2004 (while having little independence from its minister). This makes the regulatory environment less certain and thus has delayed investment (PricewaterhouseCoopers (2004)). And a non-contestable, ad hoc, and non-transparent government underwrite of gas-based new generation risks was granted in 2004 to a state-owned generator, raising fears that competing generation proposals would risk becoming unviable (not least wind-farm schemes, which had been made viable by subsidies). Even the new chairman of the Electricity Commission was critical of this move, stating “The country needs more generation but this is the wrong way to get it. The new deal risks driving new private investment out of the electricity sector, and we need that investment. Investors are keen to know the Government will not intervene on the side of the state-owned enterprises, and the Government has not provided that assurance.”²¹ Indeed, Marsden et al. (2004) argue that government, as owner of state-owned generating companies, may be accepting lower rates of return on their investments than would its private sector counterparts.

The combined effect of these factors should be predicted to make private generation investment less viable, and thus tilts the responsibility for new investment and adequacy back to government. In effect, government’s assumption of responsibility for adequacy, and resort to regulation to (try to) achieve this, begins the process of ensnaring it in the regulatory “tar baby” while simultaneously driving off private investors who could help it secure that objective (and share responsibility for any failures to do so). Under this course, future supply insecurity will become, as it was pre-liberalization, a clear manifestation of government and regulatory failure, not of any shortcomings inherent in electricity energy markets.

6. A STABLE STRATEGY TO SUPPORT ELECTRICITY INVESTMENT?

6.1. General Approaches

There are at least two ways to approach the question of security of supply/adequacy. One attempts to resolve any real or perceived deficiencies in market mechanisms by refinements or additions that lessen their impact and/or enhance market functions. The other takes those deficiencies as given, and instead looks to other means for their mitigation, such as altogether new mechanisms, with the potential to either enhance or frustrate market functions. In the present case it will be argued that the former approach is more likely to lead to self-reinforcing and stable solutions, whereas the latter has features which are self-defeating and are therefore inherently unstable. In this section measures are suggested as a self-reinforcing strategy to support electricity investment and enable adequacy. They include the partial accommodation and neutering of market power, in part by means of encouraging greater demand-side responsiveness. They also include vertical integration as a structural means to encourage investment, being an alternative to long-term contracting. And they suggest a regulatory strategy more likely to secure the desired outcomes.

²⁰ *After Aqua: NZ’s electricity future*. Address by Minister of Energy, Hon. Pete Hodgson, National Power NZ 2004 conference, Auckland, March 2004.

²¹ *Power station sparks first rift*. Dominion Post, Wellington, 13 August 2004.

6.2. Market Power and Price Caps

An initial hurdle that must be surmounted when advocating market-based approaches to adequacy provision is that of market power, and price caps instituted to moderate its effect (particularly when it manifests during shortage- or failure-induced price spikes). It would appear to be a given that oligopolistic generation is the norm in liberalizing electricity systems. Unfettered competition remains an ideal, and one which is unlikely to be attained in practice given current technologies (although advances in small-scale generation and superconductor-based energy storage and transmission have the potential to fundamentally change this). As such, it is important to be realistic regarding the price performance expected of imperfectly competitive generation markets. A zero-tolerance approach is possibly too onerous, particularly where it has the effect of destroying investment signals due to low and/or arbitrary price caps. Indeed, Bidwell and Henney (2004) argue that price spikes in energy-only markets are an unreliable means of funding generation investment because of their unpredictability and political unsustainability, and go so far as to suggest that market power exploitation is necessary (absent capacity mechanisms), although untenable, to provide the requisite returns. Roques et al. (2004) also note the potential of horizontal integration to enhance market power and thereby improve investment viability.

In fact market power can not only better fund investment, but possibly encourage strategic overinvestment as a means to deter competitive entry. Evans et al. (2003) note the trade-off between static and dynamic efficiency, with competition maximising static efficiency at the expense of innovation and investment, and market power encouraging greater investment and thus dynamic efficiency at the expense of static efficiency. Athey and Schmutzler (2001) show how investment can be used as an anti-competitive tool, which illustrates a flipside risk of capacity markets – creating excess capacity that can be used to entrench incumbent dominance. While Botterud and Korpas (2004) find that the polar extreme of monopolistic generation would in fact lead to investment deferral, perhaps suggesting capacity mechanisms, Creti and Fabra (2003) on the other hand find that the conditions supporting the optimal introduction of capacity mechanisms are more stringent under the most competitive scenarios.

Ironically, the introduction of price caps to curb market power can necessitate other measures to encourage investment to ensure adequacy, but may then do so at the risk of enhancing market power. Furthermore, Stoft (2002, p. 163) notes that setting price caps equal to VOLL in fact “provides strong incentives for the exercise of market power.” Hence by allowing market power a measure of free-rein (e.g. higher or no price caps), or implementing capacity mechanisms to elicit investment otherwise choked off by price caps, generation investment can be stimulated and the prospects of adequacy enhanced. The regulatory trade-off this induces is noted by Roques et al. (2004), and in New Zealand it is indirectly accommodated by public benefit tests when the competition watchdog assesses mergers or acquisitions leading to lessened competition (Commerce Commission (2004)). The challenge is to ensure that competition and electricity sector regulators are mindful of these trade-offs and properly specify their objective functions to enhance overall social welfare. If encouraging adequacy through viable investment is considered a welfare-enhancing priority, it may be necessary to let the beast feed.

6.3. Demand-Side Initiatives

The extent of market power and its social costs, however, need not be taken as given. It is generally accepted that greater demand-side price-responsiveness would diminish the impact of any market power exploitation, especially during price spikes (e.g. Roques et al. (2004)). If electricity demand becomes more price-elastic, then, for example, the rewards to strategic withdrawal or withholding of capacity by generators with market power are reduced. This not only diminishes the need for investment- and adequacy-reducing price caps, and leads to lower energy price volatility, but inherently relieves the pressures giving rise to

inadequacy. The more demand is shed in response to rising prices during supply shortages, the less is the generation capacity needed to meet that demand – interruptible load is relatively cheaply substitutable for new capacity. As Roques et al. (2004, p. 28) put it, “Improved demand participation allowing for load-shedding in times of scarcity would be much less costly than a gold plated electricity system.” Certainly where wholesale electricity markets lack demand-side bidding, an obvious first-step to improve market arrangements would be to introduce such bidding.

The reluctance of consumers to adopt real-time pricing and apparent dislike for electricity price risk was noted in the earlier discussion regarding consumers’ willingness to pay for supply security. Indeed, even large customers on real-time plans and with the capacity and incentives to adopt more energy-efficient technologies and flexible consumption practices tend to hedge their price exposure via financial hedge contracts. And smaller customers, by design or by accident, typically face only monthly billing with often single-price plans. However, many customers, large and small, do in fact contract to provide system operators or suppliers with interruptible load. Even smaller customers are offered rebates to reduce their demand relative to average during sustained shortages despite not having real-time meters. Increasingly power exchanges offer successively smaller customer classes the opportunity to trade surplus energy for profit. Indeed, during price spikes large customers with hedge contracts reschedule maintenance and production so that energy savings can be sold at high spot prices for substantial profit. The challenge of encouraging greater demand-side responsiveness is not so much about exposing all customers to energy prices, which tends to encourage defensive hedging by customers unaccustomed to bearing price risk, but rather providing customers with opportunities to profit by changing their consumption behaviour. To improve adequacy such measures can be supported by steps identified earlier, such as capacity subscriptions with load-limiting fuses, pricing plans that reward suppliers’ options to interrupt load via ripple control on electric water heating (as in New Zealand), or insurance or statutory guarantees offering customers price-based choices about their preferred level of supply security (which gives suppliers strong incentives to manage the risks that then shift onto them). These measures signal consumers’ willingness to pay for security of supply, the absence of such signals being identified by Hobbs et al. (2001) and Bidwell and Henney (2004) as the other key reason (besides the political untenability of price spikes in the presence of market power) why energy only electricity markets will underprovide adequacy. Rather than treating demand-side inadequacies as given and proposing complex mechanisms to compensate for those inadequacies, these measures aim to fix any problems at their source.

6.4. Vertical Integration

The use of long-term contracts between generators and purchasers is widely regarded as a way to mitigate market power and encourage generation investment (e.g. Stoff (2002)). By committing generators’ output at fixed prices this reduces their incentive to manipulate electricity prices where they have an ability to do so. Simultaneously such contracts provide generators with greater certainty as to longer-term investment returns than that achievable through volatile spot-price energy sales, thus supporting investment. The general illiquidity of forward contract markets beyond one year is a common complaint, however, and legislation was recently passed in New Zealand empowering the industry regulator to impose forward sales by generators (and forward purchases by wholesale buyers) under such contracts.

The recent New Zealand legislation and general concern regarding forward contract illiquidity is somewhat misplaced given the obvious alternative available to achieve similar, indeed better, outcomes. Most liberalizing markets began as vertically and horizontally integrated monopolies, and quickly embarked on partial or full de-integration of generation, transmission, distribution and/or energy retailing. Spain and New Zealand are unusual in this respect, in that almost complete vertical integration of generation and retailing is currently the norm. In New Zealand vertical integration was a rapid innovation, since for decades

previously distribution had been combined with retailing, and separately owned from generation and transmission (local government owning distribution/retailing, central government generation/transmission). When the state-owned generator was finally separated into competing companies in April 1999, restrictions on vertical integration were also lifted. At the same time legislation came into force requiring owners of lines companies to divest their ownership of either those lines activities, or any other competitive activities such as generation or energy retailing. The result was a rapid structural realignment of the sector, with energy retailing businesses being quickly acquired by competing generators (hence retail competition continued with this realignment). That this realignment should occur endogenously rather than by government fiat is an important indication that it serves some innate economic purpose.

When energy retailing is owned separately from generation, this places great importance on the ability of retailers to manage their exposure to price risk on their net purchases. Long-term supply contracts are in principle an important means for them to do so, but experience shows this is not a reliable approach (and given the risk-repackaging similarities between retail banks and retail energy companies, it must be asked why energy retailers should rely only on wholesale purchases and not embedded generation to manage their supply costs anyway). Failure to properly manage wholesale electricity price exposure exposes non-integrated energy retailers to the risk of financial collapse. As discussed in Meade (2001), certain characteristics of electricity prices mean that the vertical integration of generation and retailing represents an economically efficient form of hedge against volatile wholesale energy prices. With only their net sales or purchases then exposed to wholesale prices, vertically integrated “gentailers” have thus replicated the desired effects of supply contracts, and with greater duration than that achievable using only traded, short-term contracts. Exposure to volatile wholesale energy prices is then replaced by the less volatile difference between retail prices (many set under fixed supply contracts) and generation costs (some of which, at least, are discretionary and controllable). This provides smoother returns to investment over time, reduces the risk of financial collapse, and facilitates the provision of a greater range of price-risk management options to consumers. New Zealand’s 1999 structural realignment “experiment” suggests that the costs of ownership are in this case less than the costs of market contracting, making ownership of energy retailing by generation a more natural organisational model than contracting between these activities (a potent illustration of Coase (1937) in action).

Vertical integration simultaneously overcomes some of the informational asymmetry problems and diseconomies of scale possibly imposed with structural separation. Like long-term supply contracting, it also diminishes the incentive for generators with market power to exploit that power, further diminishing the rationale for price caps and hence compensatory capacity mechanisms. Kuhn and Machado (2004) suggest that the absence of major concerns in Spain about excessive energy prices is most likely a consequence of vertical integration, since net demand positions are what matter for such firms, and so long as their market power is similar, prices may not differ much from competitive levels. Evidence by Mansur (2003) from PJM in the US, which also has vertically integrated firms, supports this view, with prices at competitive levels except for firms that are large net sellers (who withhold output relative to competitive levels). Furthermore, Benz et al. (2003) show that investment incentives are improved under vertical integration as compared with de-integration (although, as noted previously, not uniformly so). Helm (1994) argues that vertical integration (like long-term supply contracts) provides a hedge against the risk of regulatory hold-up where investments are sunk and regulatory horizons are shorter than investment lives. For all these reasons vertical integration contributes to improved investment incentives, compensating for any market deficiencies otherwise leading to underinvestment in generation, and thereby naturally contributes to the achievement of adequacy. It perhaps provides a feasible and superior successor to the merchant generation model said by de Luze (2003) to have been “destroyed” in 2002.

6.5. Regulatory Approach

Political intervention and regulation pose ever-present risks to market-directed investment, as predicted by the theory of public choice and regulatory economics (e.g. Stigler (1971), Joskow and Rose (1989), Helm (1994), Baarsma et al. (2004)). While regulation has traditionally been considered a benign process seeking to remedy market or other failings in some socially optimal manner, more recently it is recognised that regulators (and the regulated) confront potentially disadvantageous information asymmetries, meaning they struggle to identify, let alone make, socially-optimal decisions. It is also now recognised that regulators' incentives will not perfectly align with those of politicians, consumers, the regulated or investors (i.e. that they pose principal-agent problems common to other governance situations), meaning their incentive to make socially-optimal decisions is also in question. Furthermore, political cycles of shorter duration than investments expose long-lived investments to hold-up from policy and regulatory changes (or overhauls). Individual voters have only very diffuse incentives to involve themselves in the political and regulatory processes, meaning that organised interest groups (e.g. environmental lobbies) can capture a disproportionate share of benefits from those processes, leading to outcomes potentially hazardous to investment, but without bearing the full costs of their actions (so-called "rent-seeking" behaviour). With electricity liberalization being an exercise in re-regulation as much as deregulation, it should come as little surprise that more than 80% of respondents in the 2004 Capgemini Global Utility Survey cited risks of regulatory or political intervention as being of major impact on their investment decisions.

It should also be noted, however, that the political and regulatory processes are also exposed to capture by business lobbies seeking to secure advantage over rivals or customers (Stigler (1971), Helm (1994)). The origins of rate of return regulation of electric utilities in the US are a pertinent case in point (Stoff (2002), Michaels (1996)). Since generation investors have much to lose from the political and regulatory processes, they also have strong incentives to capture those processes. Some will be better able to do so than others, even within a given industry, meaning that capture by business interests may favour some but prejudice others both within and across industries, unequally affecting their investment prospects, and with the risk of reversals should currently favoured parties fall out with future governments or regulators.

Debate continues regarding the extent to which investors can hedge their regulatory risks (Roques et al. (2004)), although the above suggestions show how some important triggers for regulation (i.e. market power, price volatility, and real or perceived underinvestment) can be relieved. It is too much to expect that future governments will regard themselves constrained by the decisions of previous ones, so the risk of future policy changes affecting the value of sunk, large and long-lived generation investments cannot be avoided (although a shift away from large and inflexible to more smaller-scale and flexible investments may mitigate this risk). What can be done, however, to moderate regulatory risk where its triggers remain, is to reduce regulatory discretion.

Helm (1994) sees excessive regulator discretion (also involving inconsistency between different regulators) as the central problem of the British regulatory system. The licencing system used in Britain is the main means of limiting this problem (Roques et al. (2004)), involving commercial contracts enforceable against government in the event of breach. These are inadequate, however, to fully mitigate the risk of regulator discretion, let alone of policy change. Roques et al. discuss the possible introduction of a "revenue standard" as an administrative approach that prevents or compensates cost recovery shortfalls resulting from new regulations. Failing that it may simply be preferable for legislators to err on the side of constraint rather than liberality when writing regulators' empowering rules. This might involve legislative undertakings not to impose policy changes or regulations, for example where they give rise to investment hold-up. It would then take specific legislation to overcome legal challenges to policy changes or regulations that do so, which fetters their imposition to a degree and requires a transparent consideration of their effects.

In the main what is required is a political attitude of restraint, to not rush to regulate for solutions and to limit regulatory discretion rather than increase it. This is naturally a major challenge given the weaknesses of political systems identified in the public choice literature and theory of regulation. It is clearly at variance with the rapid change of approach evident in the New Zealand electricity sector of the past five years (Meade (2004), Evans and Meade (2005)), with an initial move towards targeted, explicit regulation spawning further expansions of regulatory discretions and options, increasing ministerial control of the sector, and snowballing risks of further regulation to compensate for the effects of earlier interventions.

Where political restraint and reduced regulatory discretion is abandoned in favour of political activism, mushrooming regulatory discretions and intervention, the natural consequences are a reduction in committed investments, crowding out and deterring of new private investment, reduced adequacy, increased electricity price volatility and market power concentration, further interventions still, and a gradual (but possibly sudden) reversal of the liberalization process (in effect if not by intent). Consumers and taxpayers once again face being forced to pay for inefficient overcapacity (which of itself does not guarantee adequacy). Government risks finding itself once again solely responsible for security of supply, and hence solely accountable for the inevitable instances when technical, weather-related, economic or other reasons mean that adequacy is not achievable. By unnerving private investors in generation, which de Luze (2003) shows to be an actual risk, this simultaneously undermines government's ability to find new ways of providing adequacy.

6.6. Summary

In summary, it is suggested that parsimonious and self-reinforcing measures to secure adequacy should be preferred to alternatives which involve both increasing and self-perpetuating layers of complexity as well as inherent obstacles to their success. If price caps are introduced to constrain generator market power, this has numerous features frustrating the achievement of adequacy. First, price caps can in fact enhance market power (Stoft (2002)), creating an unhelpful reinforcement for the intervention. Capped scarcity rents, risks of further regulatory intervention and where caps are effective, reduced market power, all serve to undermine generation investment. The introduction of capacity mechanisms to compensate for this can enhance market power by increasing anti-competitive excess capacity, creating another unhelpful feedback mechanism. Such mechanisms also dampen investment signals and over-stimulate consumption by distorting energy prices, which in fact serves to undermine adequacy. This in turn raises further risks of regulatory intervention and undermines private investment (perhaps in favour of government investment, with likely associated inefficiencies and overcapacity), making the achievement of adequacy even more uncertain. It is hard enough for regulators and politicians suffering informational asymmetries and problematic incentives to implement simple interventions in non-distortionary ways; how much harder complex, multi-tiered and interrelated interventions?

The alternative suggested is to focus on inducing (not imposing) demand-side responsiveness, and allowing vertical integration of generation and retailing. This reduces the need for new generation to achieve adequacy (thus making it easier to achieve), and diminishes market power (thus mitigating the need for interventions such as price caps and the associated risk of unanticipated and/or arbitrary cap changes). It therefore also reduces the need for capacity mechanisms. In so doing it avoids the unhelpful feedback mechanisms created by price caps and capacity mechanisms. It reduces the need for these measures in the first place rather than taking their root causes as given, and avoids imposing solutions that can exacerbate those causes and require self-perpetuating and ever-increasing interventions.

7. CONCLUSIONS

Debate rages regarding the efficacy and desirability of interventions designed to overcome real or perceived risks that electricity markets will fail to elicit the generation investments required to ensure security of electricity supply. For some, the rationale for capacity mechanisms is thought to lie in the immaturity of electricity markets which would otherwise be expected to perform as desired, and presents a justification for transitional interventions to compensate for current market shortcomings. The obvious counter to this view is that such interventions create artificial obstacles to the achievement of market maturity, being substitutes for market refinements and learning. Others present capacity mechanisms as a pragmatic response to the risk of government intervention where ubiquitous generator market power is feared to cause undue energy price volatility and heights. Indeed, where there are government interventions to curb market power, such as price caps, capacity mechanisms are then arguably a necessary imposition to achieve the level of generation investment required for security of supply.

However, as Selgin (2003, p. 3) puts it in relation to bank sector wholesale payments, which exhibit technical features and systemic risks akin to those in electricity networks, “the literature on wholesale payments has tended to confuse regulatory failure with market failure, thereby diverting attention from potential first-best solutions to alternatives that may not even qualify as second best.” His argument parallels those here, that market failures relating to electricity externalities with potentially catastrophic consequences are based not on empirical evidence but more on references to inadequate benchmarks (e.g. electricity markets with zero price elasticity). There is a risk that the clamour to implement mechanisms to attempt to solve electricity supply security problems that may not in fact exist, and which are a consequence not of electricity markets but rather tractable features or artificial impositions, will frustrate the achievement of the desired adequacy. Much of the impetus for interventions to overcome perceived market shortcomings has come from either a need to compensate for other interventions, misunderstandings as to the nature of adequacy (e.g. that it is a public good, or that volatile electricity prices mean it is not being achieved), unquantified albeit plausible concerns about externalities, or a failure to recognise that those interventions are likely to be at least as vulnerable to any shortcomings attributed to electricity markets. To the extent a problem with electricity markets exists, the challenge is to identify strategies that resolve the problem at a fundamental level and in a self-reinforcing way, and to avoid those which can exacerbate the fundamental problem, require ever-expanding intervention, and contain elements that are self-defeating.

The approach advocated here relies less on capacity mechanisms to compensate for investment deficiencies due to price caps or institutional shortcomings, and more on measures to improve fundamental investment incentives while reducing problems such as market power which otherwise induce intervention. In the first instance this requires improvements in institutional arrangements on the demand-side, but also the permitting of structural solutions such as endogenous vertical integration of generation and energy retailing. It also requires an acceptance that market power can contribute to the achievement of adequacy and hence that regulatory tradeoffs are required when mitigating market power (which from the perspective of competition watchdogs may be less of an issue than for electricity industry regulators). More challenging is the suggestion of political restraint and reduced regulatory discretion, which should be facilitated by such improvements, but which also requires that naturally strong temptations to intervene are resisted. At its heart it suggests that confidence be had in market mechanisms that prove adequate in numerous other settings, which confidence lies at the heart of the liberalization process anyway, and that adherence to the centralized, control-based engineering model that has dominated electricity systems for much of their history be critically regarded. If this cannot be achieved, and unnecessary or counter-productive interventions imposed, the risk is that market mechanisms will never be able to fully perform their function, and the liberalization process will thus be undermined and ultimately reversed. The consequence of that would be to return

responsibility for investment and adequacy to government, at the expense of consumers and taxpayers, and with no greater guarantee of success.

8. REFERENCES

1. Athey S., Schmutzler A.: *Investment and market dominance*. RAND Journal of economics, vol. 32, no. 1, Spring 2001, pp. 1-26.
2. Baarsma B., Poort J., Teulings C., de Nooij M.: *Private investments in new infrastructures*. SEO Amsterdam Economics, www.seo.nl.
3. Benz M-A., Buehler S., Schmutzler A.: *Infrastructure quality in deregulated industries: is there an underinvestment problem?* CEPR Discussion Paper, no. 3836, March 2003.
4. Bidwell M., Henney A.: *Will Neta ensure generation adequacy?* Power UK, Issue 122, April 2004, pp. 10-26.
5. Boiteux M.P.: *La tarification des demandes en pointe: application de la theorie de la vente au cout marginal*. Revue Generale de l'Electricite, 1949.
6. Borenstein S.: *The trouble with electricity markets: understanding California's restructuring disaster*. Journal of Economic Perspectives, Winter 2002, pp. 191-211.
7. Botterud A.: *Generation investments in restructured power systems: perspectives from the Nordic power market (Nord Pool)*. HEPG, October 2004.
8. Botterud A., Korpas M.: *Modelling of power generation investment under uncertainty in liberalised electricity markets*. 2004.
9. Caramanis M.: *Investment decisions and long-term planning under electricity spot pricing*. IEEE Transactions on Power Apparatus and Systems, vol. PAS-101, no. 12, December 1982, pp. 4640-4648.
10. Chao H-P.: *Peak load pricing and capacity planning with demand and supply uncertainty*. The Bell Journal of Economics, vol. 14, 1983, pp. 179-190.
11. Coase R.H.: *The Nature of the Firm*. Economica, n.s., vol. 4, November 1937.
12. Coase R.H.: *The lighthouse in economics*. The Journal of Law and Economics, vol. 17, no. 2, October 1974, pp. 357-376.
13. Coase R.H.: *The firm the market and the law*. University of Chicago Press, 1988.
14. Colegrave F., Denne T., Hale R., Small J., Twomey I.: *Oil security*. Report prepared for Ministry of Economic Development, www.med.govt.nz, November 2004.
15. Commerce Commission: *Mergers and acquisitions guidelines*, www.comcom.govt.nz, January 2004.
16. Creti A., Fabra N.: *Capacity markets for electricity*. November 2003.
17. de Luze G.: *Investment in power generation: a banker's perspective*. IEA/NEA Workshop on Power Generation Investment in Liberalised Electricity Markets, Paris, March 2003.
18. de Nooij M., Koopmans C., Byvoet C.: *The demand for supply security*. The European Electricity Market EEM-04: Challenge of the Unification, September 2004, pp. 77-87.
19. de Vries L.J., Hakvoort R.A.: *The question of generation adequacy in liberalised electricity markets*. INDES Working Papers, Centre for European Policy Studies, No. 5, March 2004.
20. Doorman G.: *Capacity subscription and security of supply in deregulated electricity markets*. SINTEF Energy Research, 2003.
21. Electricity Commission: *Cover paper for "Analysis of the state of competition and investment and entry barriers to New Zealand's wholesale and retail electricity markets"* by Murray K., Stevenson T., August 2004.
22. Electricity Corporation of New Zealand: *Electricity pricing in New Zealand*. October 1991.
23. Electricity Shortage Review Committee: *The Electricity Shortage 1992*, 1992.
24. Evans L., Meade R.: *Economic Analysis of Financial Transmission Rights (FTRs) With Specific Reference to the Transpower Proposal for New Zealand*. New Zealand Institute for the Study of Competition and Regulation Inc, www.iscr.org.nz, September 2001.
25. Evans L., Meade R.: *Contemporary electricity reform in New Zealand*, New Zealand Institute for the Study of Competition and Regulation Inc., forthcoming 2005.

26. Evans L., Quigley N., Zhang J.: *Dynamic efficiency, regulation and price control*. Canadian Journal of Economics, vol. 36, no. 2, 2003, pp. 463-474.
27. Ford A.: *Cycles in competitive electricity markets: a simulation study of the western United States*. Energy Policy, vol. 27, 1999, pp. 637-658.
28. Fraser H., Lo Passo F.: *Developing a capacity payment mechanism in Italy*. The Electricity Journal, November 2003, pp. 54-58.
29. Friedman D.: *Private and political markets both fail: a cautionary tale about government intervention*. New Zealand Business Roundtable, December 2004.
30. Galvin, B.: *Review of Electricity Planning and Electricity Generation Costs*. The Treasury, Wellington, 1985.
31. Helm D.: *British utility regulation: theory, practice, and reform*. Oxford Review of Economic Policy, vol. 10, no. 3, 1994, pp. 17-39.
32. Hirst E., Hadley S.: *Maintaining generation adequacy in a restructuring U.S. electricity industry*. ORNL/CON-472, October 1999.
33. Hobbs B.F., Inon J., Stoft S.E.: *Installed capacity requirements and price caps: oil on the water, or fuel on the fire?* The Electricity Journal, July 2001, pp. 23-34.
34. International Energy Agency: *Security of supply in electricity markets: evidence and policy issues*. OECD, Paris, 2002.
35. Jaffe A.B., Felder F.A.: *Should electricity markets have a capacity requirement? If so, how should it be priced?* The Electricity Journal, December 1996, pp. 52-60.
36. Joskow P.: *Contributions to the theory of marginal cost pricing*. Bell Journal of Economics, vol. 7, Spring 1976, pp. 197-206.
37. Joskow P., Rose N.L.: *The effects of economic regulation*. in Schmalensee, R., Willig R.D. (eds), Handbook of Industrial Organization, vol. II, Elsevier Science Publishers, 1989.
38. Joskow P., Tirole J.: *Reliability and competitive electricity markets*. NBER Working Paper, no. 10472, May 2004.
39. Kuhn K-U., Machado M.P.: *Bilateral market power and vertical integration in the Spanish electricity spot market*. CEMFI Working Paper, no. 0414, September 2004.
40. Laffont J-J., Tirole J.: *The politics of government decision-making: a theory of regulatory capture*. The Quarterly Journal of Economics, vol. 106, no. 4, November 1991, pp. 1089-1127.
41. Mansur E.: *Vertical integration in restructured electricity markets: measuring market efficiency and firm conduct*. October 2003.
42. Marsden A., Poskitt R., Small J.: *Investment in the New Zealand electricity industry*. UniServices, October 2004.
43. Meade R.: *Options, hedges and gamma in the electricity business*. Competition & Regulation Times, Issue 6, New Zealand Institute for the Study of Competition and Regulation Inc., www.iscr.org.nz, November 2001, pp. 11-12.
44. Meade R.: *Decentralization and re-centralization of electricity industry governance in New Zealand*. The European Electricity Market EEM-04: Challenge of the Unification, September 2004, pp. 57-65.
45. Michaels, R.J.: *Stranded investments, stranded intellectuals*. Regulation, vol. 19, no. 1, 1996.
46. Ministry of Economic Development: *New Zealand energy outlook to 2025*. www.med.govt.nz, October 2003.
47. Ministry of Economic Development: *Government policy statement on electricity governance*. www.med.govt.nz, October 2004.
48. Oren S.: *Capacity payments and supply adequacy in competitive electricity markets*. VII Symposium of Specialists in Electrical Operational and Expansion Planning, Brazil, May 2000.
49. Ostrom E.: *Collective action and the evolution of social norms*. Journal of Economic Perspectives, vol. 14, no. 3, Summer 2000, pp. 137-158.
50. PricewaterhouseCoopers: *Infrastructure stocktake: Infrastructure audit*. www.med.govt.nz, January 2004.

51. Rochlin C.: *Resource adequacy requirement, reserve margin, and the public goods argument*. The Electricity Journal, April 2004, pp. 52-59.
52. Roques F., Newbery D.M., Nuttall W.J.: *Generation adequacy and investment incentives in Britain: from the pool to NETA*. CMI Working Paper 58, October 2004.
53. Selgin G.: *Wholesale payments: questioning the market-failure hypothesis*. September 2003.
54. Sinclair Knight Merz: *Power generation options for New Zealand*. Report prepared for Ministry of Economic Development, www.med.govt.nz, July 2003.
55. Small J.P.: *The timing and scale of investment under uncertainty*. CNREC, University of Auckland, June 1999.
56. Stigler G.J.: *The theory of economic regulation*. The Bell Journal of Economics and Management Science, vol. 2, no. 1, 1971, pp. 3-21.
57. Stoft S.: *Power System Economics: Designing Markets for Electricity*, Wiley Inter-Science, 2002.
58. Teece, D.: *Economies of scope and the scope of the enterprise*. Journal of Economic Behavior and Organization, vol. 1, 1980, pp. 223 – 247.
59. Vazquez C., Rivier M., Perez-Arriaga I.J.: *A market approach to long-term security of supply*. IEEE Transactions on Power Systems, vol. 17, no. 2, May 2002, pp. 349-357.